

## The Packing of $\alpha$ -Helices: Simple Coiled-Coils

BY F. H. C. CRICK

*The Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems,  
The Cavendish Laboratory, Cambridge, England*

(Received 14 March 1953)

It is shown that when  $\alpha$ -helices of the same sense pack together they will probably do so about  $20^\circ$  away from parallel. For very long chains this may lead to a coiled-coil. The two simplest models—the two-strand rope and the three-strand rope—are described, and used to illustrate the diffraction theory already developed. It is shown that they would give a diffuse  $\alpha$ -pattern. Possible examples of these models are briefly discussed.

### Introduction

It is now firmly established that the structure of the synthetic polypeptide polymethylglutamate is based on the  $\alpha$ -helix of Pauling, Corey & Branson (1951). The general similarity of the X-ray diffraction pattern of all synthetic polypeptides in the  $\alpha$  form so far examined makes it very probable that they are all based on this helix.

Pauling & Corey (1951) suggested that the  $\alpha$ -helix could also explain the ' $\alpha$ -keratin' diffraction pattern, which the pioneer work of Astbury and his school has shown to be given in various degrees of perfection by porcupine quill, hair, horn, muscle, epidermin, fibrinogen and related proteins. (The k-m-e-f group).

The main characteristics of this diffraction pattern are meridional arcs at spacings of about 5.15 and 1.5 Å, and a group of reflexions on and near the equator at spacings around 10 Å (McArthur, 1943; Perutz, 1951). The two main difficulties in fitting the  $\alpha$ -helix to this structure are:

(1) The 5.15 Å reflexion on or very close to the meridian of the reciprocal lattice. A straight perfect  $\alpha$ -helix, parallel to the fibre axis, gives a strong layer-line at a spacing of 5.4 Å, but the intensity on the actual axis of the reciprocal lattice should be zero. This argument is less precise if the arrangement of the side-chains is not strictly helical.

(2) The density. If the centre of the broad equatorial reflexion at 9.8 Å is taken as the (10.0) reflexion of a simple hexagonal lattice the calculated density for  $\alpha$ -keratin is too low. (It is not clear that all members of Astbury's k-m-e-f group have a high density.)

To explain the 5.15 reflexion on the meridian Crick (1952) suggested that the  $\alpha$ -helix might be deformed into a coiled-coil. It was shown that the energy involved in this deformation was likely to be small. The reason suggested for the deformation was that the non-integer nature of the  $\alpha$ -helix made it more likely that two helices having the same sense of twist would pack together at an angle rather than exactly parallel, and that this would lead to a coiled-coil.

Simultaneously Pauling & Corey (1953) put forward a detailed model for  $\alpha$ -keratin based on coiled-coils. They suggested that the origin of the deformation was a repeating sequence of amino acids, a repeat every seventh residue being required for two-thirds of the  $\alpha$ -helices in the structure, and a repeat every fourth residue for two-ninths of them.

This model can explain the simultaneous existence of both the 5.15 and the 1.5 Å reflexions on the meridian. It also broadly explains the equatorial reflexions, and probably the near equatorial reflexions of African porcupine quill. The density calculated for this model is nearer the observed value, but it is still on the low side.

Although the fibrous  $\alpha$ -proteins all give the  $\alpha$ -keratin pattern in the loose sense, they give it to varying degrees of perfection, African porcupine quill having sharp spots (McArthur, 1943) whereas films of tropomyosin (Astbury, Reed & Spark, 1948) give only a diffuse pattern. The more diffuse the  $\alpha$ -keratin pattern becomes, the more it resembles a very diffuse pattern from a synthetic polypeptide, that is, a diffuse  $\alpha$ -helix pattern. It seems unlikely to me that the diffuse  $\alpha$ -keratin patterns are all disordered versions of exactly the same structure. It is more probable that their similarity reflects a general type of structure, which I believe to be  $\alpha$ -helices arranged in a non-parallel manner for reasons of packing. It is unlikely that tropomyosin for example, which may well be only two  $\alpha$ -helices wide, has the detailed structure proposed by Pauling & Corey, which has been based on data from African porcupine quill.

In a previous paper (Crick, 1953) the general formula for the X-ray scattering from a set of atoms arranged on a coiled-coil has been derived. In the present paper the simple packing argument already put forward (Crick, 1952) is explained more fully. It is based in the first instance on a model in which the side-chains are idealized into uniform knobs which are then made to pack into the holes (the spaces between knobs) of the adjacent helix.

The two simplest coiled-coil structures—the two-strand rope and the three-strand rope—are described, and used to illustrate the diffraction theory. It is shown that these simple structures will give a diffuse ‘ $\alpha$ -keratin’ pattern, and may be the basis of long thin molecules like tropomyosin.

### Types of $\alpha$ -helices

Although the amino acids making up the polypeptide chain are all L-amino acids it is possible to build models of the  $\alpha$ -helix in two different ways, depending upon whether the polypeptide chain follows a right-handed or a left-handed helix. This difference is sometimes expressed by saying that the  $\beta$  carbon atoms may be in position 1 or position 2 (Pauling & Corey, 1951). If we have a model with  $\beta$ C in position 1, and then change it over to position 2, without any other alteration, we shall change the  $\alpha$  carbon atom from an L to a D configuration. If we now look at this new structure in a mirror, the image we see there will have an L configuration for the  $\alpha$  carbon atom once again. The mirror will also change the sense of twist of the helix. Thus the two possible structures are the one we built first, and the mirror image of the second.

If we turn a right-handed  $\alpha$ -helix upside down, it stays a right-handed helix, but the direction of the polypeptide chain is reversed in space. The sequence C-N-C is like an arrow fixed to the helix to show its direction.

We thus have four possibilities for an  $\alpha$ -helix whose axis is parallel to some given direction in space, since

it must have both its sense of twist and its direction given to fix it unambiguously, apart from the possibility of rotating it about its own axis. These four possibilities are illustrated diagrammatically in Fig. 1.

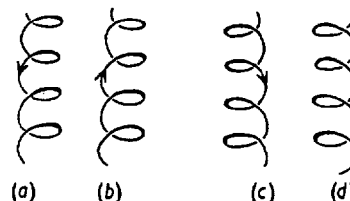


Fig. 1. (a) and (b) Right-handed helices. (c) and (d) Left-handed helices. The arrow symbolizes the direction of the sequence of atoms in the polypeptide chain.

It is possible that all the  $\alpha$ -helices found in polypeptides and proteins have the same sense. Yakel, Pauling & Corey (1952) have claimed that for polymethylglutamate the  $\beta$ C2 position is probably established by the data. Riley & Arndt (1952) favoured the same position for bovine serum albumin from their radial-distribution studies. On the other hand, more recently (Riley & Arndt, 1953) they suggest that insulin may have both types of chain.

In none of these cases, however, can the matter be said to be definitely established, and the question must still be regarded as open.

I shall discuss here mainly the packing together of  $\alpha$ -helices of the *same* sense. It is easy to see that the packing of helices of opposite sense is simpler, since their side-chains can mesh together in the way that two gear wheels revolving in opposite directions can interlock. They would, therefore, be expected to pack parallel to each other. It will be shown that it is probably difficult for two  $\alpha$ -helices of the same sense to do this.

### The basic packing scheme

Consider, therefore, the packing together of two  $\alpha$ -helices of the same sense. For the moment we will also assume that they have the same direction. We shall idealize the side-chains into knobs of uniform size and shape, without specifying too closely what this is, and examine the pattern they make in space.

Imagine that we have a model of such an  $\alpha$ -helix. Let us pretend that we wrap round it, at a radius of about 5 Å, a cylinder of paper, and mark on the paper the point where each side-chain comes. We then open up the paper until it is flat again, and examine the pattern we have drawn. For an  $\alpha$ -helix having 3.6 residues per turn we shall find something like Fig. 2.

It can be seen that we have a regular pattern of points, about 7 or 8 Å apart, which slants across the paper. The broken line in Fig. 2 is not parallel to the helical axis because the  $\alpha$ -helix is a non-integer helix, and repeats nearly, but not exactly after two turns.

Let us do the same thing for a second  $\alpha$ -helix. Naturally we get the same pattern as before. If we

now try to pack our two helices together, the two pieces of paper will come into contact, and the condition that the knobs of one helix fall *between* the knobs

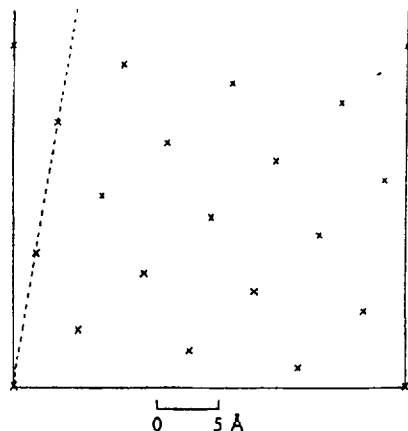


Fig. 2. The pattern formed by the side-chains of an  $\alpha$ -helix (see text). Helical axis vertical.

of the other is equivalent to saying that where the two pieces of paper are in contact—or nearly in contact—the marks on one must fall between the marks on the other. If we do this with the two bits of paper opened out flat, however, we must turn one of them over before we superimpose them, since the ‘outside’ of both sheets must come into contact.

It is immediately found that we can only do this in one way if the chains are to be approximately parallel. This is illustrated in Fig. 3. It can be seen that the

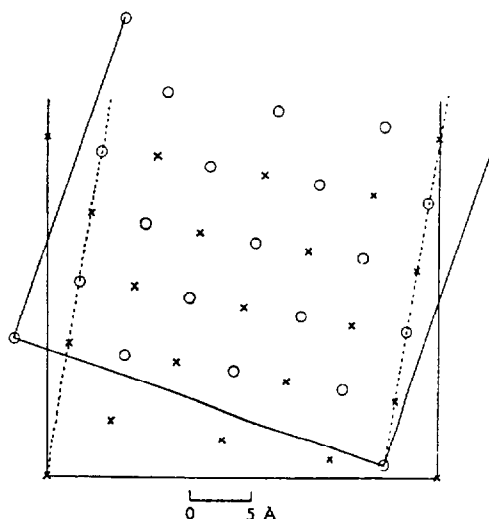


Fig. 3. The pattern formed by the side-chains of two  $\alpha$ -helices. The side-chains of one are denoted by crosses, and of the other by circles.

knobs on one helix (marked as crosses) fall nicely between the knobs of the second (marked as circles) in a systematic manner. Moreover, the patterns fit not only near the line of contact, shown as broken lines,

but right across the paper, showing that we could roll one helix round the other and they would always interlock. The reason for this is that the broken line in Fig. 2 is roughly a mirror plane in the pattern.

Each mark on Fig. 3 is surrounded by four nearest neighbours about 5 Å or so away on the paper. The exact distance in space will depend on the radius at which we put our knobs. Since the van der Waals diameter of a methyl group is 4 Å, we see that there will probably be a reasonable amount of room.

If the reader wishes to have a graphic demonstration of this form of packing he should make himself an idealized model of two short sections of  $\alpha$ -helix, using (say) the handle of a broomstick (to represent the polypeptide backbone) and knobs of plasticine (to represent side-chains) placed over nails carefully inserted in the correct positions on the broomstick. It will be found that such models pack naturally together in the manner described.

To achieve this superposition it is necessary to incline the axes of the helices to each other. The angle between them is about 20°. If the two helices remain straight, they will separate, but if we deform them slightly we can make them remain in contact indefinitely. They will then slowly wind round each other like the two wires of a piece of twin lighting cable. Every seventh knob of one will fall into every seventh hole in the other, and their axes will each follow a larger, more gradual helix with a pitch angle of about 10°.

It can be shown (Crick, 1952) that the energy to deform each helix into a curved helix or ‘coiled-coil’ is probably very small.

If the two helices have the same sense, but run in opposite directions, then we should get the same patterns on our pieces of paper as before, and thus we should again expect a similar coiled-coil, though its diffraction pattern may not be identical with that of the coiled-coil described first, because of the difference in the directions of the polypeptide backbones.

If we do not require the two helices to be approximately parallel we can find a further position in which the side-chains crudely interlock. In this pack the helices are about 70° away from parallel. This case will not be considered further here.

As mentioned earlier, if we pack together two helices of *opposite* sense our bits of paper will superpose accurately when the chains are parallel. We shall not consider this case further here either, since its diffraction pattern can be worked out from the simple theory for a straight  $\alpha$ -helix.

### The Fourier transform of a coiled-coil

The Fourier transform of a set of scattering points (atoms) arranged at regular intervals along the locus of a coiled-coil has been given by Crick (1953). The method of using this to calculate the structure factors

of a real coiled-coil structure is explained in the introduction to that paper.

The gradual helix followed by the axis of the curved  $\alpha$ -helix is called the major helix. It has radius  $r_0$ , a repeat distance of  $P$  in the  $z$  direction, and its pitch angle  $\alpha$  is given by  $\tan \alpha = 2\pi r_0/P$ . The helix which approximates to the  $\alpha$ -helix itself is called the minor helix.  $r_1$  is the distance of an atom from the axis of the minor helix. The complete structure is assumed to repeat after a distance  $c$  in the  $z$  direction, the major helix making  $N_0$  turns, and the minor helix  $N_1$  turns 'in its own frame' in that distance. The precise definition of  $N_1$  is given in the original paper. There are  $M$  atoms, equally spaced along the coiled-coil locus, in the repeat distance  $c$ . The cylindrical co-ordinate of reciprocal space are denoted by  $R, \psi$  and  $Z$ .

The Fourier transform  $C(R, \psi, Z)$  is non-zero only on the layer-lines (we shall call  $l$  the number of the layer-line) since the structure is periodic in the  $z$  direction (the fibre axis) but non-periodic in the other two directions. From equations (10) and (12) of Crick (1953) we write down the result for a right-handed major helix and a left-handed minor helix. This is

$$C(R, \psi, Z) = C(R, \psi, l/c) \\ = \sum_p \sum_q \sum_s J_p(2\pi R r_0) J_q(2\pi R r_1) J_s(2\pi(l/c)r_1 \sin \alpha) \\ \times \exp i[p(\frac{1}{2}\pi + \psi) + q(\frac{1}{2}\pi - \psi) + s\pi], \quad (1)$$

subject to the restriction that one includes only terms for which

$$N_0 p + (N_1 - N_0) q + N_1 s = l + m' M, \quad (2)$$

where  $p, q, s$ , and  $m'$  can take any integer values, positive or negative.

For the reasons explained in the original paper, namely that  $J_n(x)$  is very small for small  $x$  and large  $n$ , we are usually able to neglect Bessel functions of high order, so that we have to consider only terms for which  $p, q$  and  $s$  are all small. For similar reasons the most important terms are likely to be those for which two of the Bessel functions are of order zero. It will become apparent later on that we can usefully describe some of these as follows:

$$\begin{array}{ll} p = q = s = m' = 0 & : \text{equatorial;} \\ q = s = m' = 0, p \text{ small} & : \text{near equatorial;} \\ p = q = m' = 0, s = 1 & : 5.1 \text{ \AA meridian;} \\ p = q = s = 0, m' = 1 & : 1.5 \text{ \AA meridian;} \\ p = s = m' = 0, q = 1 & : 5 \text{ \AA near meridian.} \end{array}$$

We shall now consider the simple cases of the two-strand rope and the three-strand rope to illustrate the general character of the diffraction theory.

### The two-strand rope

To make the model precise we have to define the relation between the two  $\alpha$ -helices making up the rope. We will restrict ourselves for the moment to two helices having the same direction.

It seems probable that they will be related by a diad parallel to the axis of the structure. That is, that the line of contact will be like the broken line in Fig. 4,

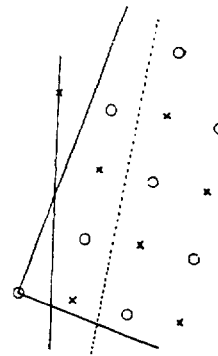


Fig. 4. Part of Fig. 3, showing an alternative position for the broken line. In the two-strand rope the broken line represents the line of contact.

rather than that in Fig. 3. Simple models suggest that in the case of the two-strand rope this arrangement gives most room for those side-chains which occur near the line of contact to turn outwards, away from the line of contact, so that packing should be easy.

For simplicity the contribution of only one of the pair of helices making up the two-strand rope will be considered in the first instance.

We take as our parameters:

$$\begin{array}{ll} \text{Turns of the major helix, } N_0 & = 1. \\ \text{Turns of the minor helix, } N_1 & = 36. \\ \text{Atoms in the set, } M & = \frac{1}{2} \times 36 \times 7 = 126. \\ \text{Repeat distance, } C & = 126 \times 1.5 \cos 10^\circ \text{ \AA} = 186 \text{ \AA}. \end{array}$$

Radius of major helix,  $r_0 = 5.2 \text{ \AA}$ , say.

Equation (2) therefore becomes

$$p + 35q + 36s = l + 126m'. \quad (3)$$

Now consider first the meridian of the reciprocal lattice, i.e.  $R = 0$ . This makes both  $(2\pi R r_0)$  and  $(2\pi R r_1)$  zero, so that, since  $J_p(0)$  is non-zero only for  $p = 0$  and since  $J_0(0) = 1$ , we obtain for the amplitude of the structure factor for a single set of atoms the value

$$C(0, 0, l/c) = \sum_s J_s(2\pi(l/c)r_1 \sin \alpha) \exp(is\pi), \quad (4)$$

subject to

$$36s = l + 126m'. \quad (5)$$

Thus since 36 and 126 are multiples of 18,  $l$  must be a multiple of 18, so that there are no reflexions on the meridian except for spacings which are sub-multiples of  $186/18 = 10.35 \text{ \AA}$ .

Such a simple result is likely to spring from an equally simple reason. Inspection of our model shows that each helix can be turned into itself (since we have tacitly assumed all side-chains are the same) by the operation of a screw axis which has a translation of

Table 1

$(l/18)$	0	1	2	3	4	5	6	7	8	9	10
$d$ (Å)	$\infty$	10.3	5.17	3.45	2.59	2.07	1.72	1.48	1.29	1.15	1.03
$s$	$\begin{cases} 0 \\ -3 \end{cases}$	-3	1	-2	2	-1	3	0	4	1	5

This is a distribution of the same type that one finds for a simple helix (Cochran, Crick & Vand, 1952).

10.35 Å (the distance apart of every 7th side-chain) and a rotation of 20°. This 20° springs from the fact that in a straight  $\alpha$ -helix, with a rotation of 100° per residue, the rotation after 7 residues is 700°. This is 20° less than 720°, which is two complete turns. The fact that  $700^\circ - 720^\circ = -20^\circ$  (not  $+20^\circ$ ) explains why the major helix has the opposite sense to the minor helix in this case. A similar argument, though from a different premiss, has been given by Pauling & Corey (1953).

The formula used by Crick (1952) for the expected meridional spacing for a coiled-coil  $\alpha$ -helix, namely  $5.4 \cos \alpha$  Å, where  $\alpha$  is the pitch angle, is seen to be incorrect. Here 5.4 Å is the axial distance per turn for the straight  $\alpha$ -helix. The above formula is correct for an *inclined* helix (angle of inclination =  $\alpha$ ), but for a coiled-coil one must also allow for the slow rotation of the major helix. Thus there is no need to postulate  $\alpha = 18^\circ$  in order to make  $\cos \alpha = 5.15/5.4$ , and in fact a tilt of half this is satisfactory if the correct formula is used.

From equation (5) we can easily write down the values of  $s$  appropriate to the various layer-lines. They are given in Table 1. High values of  $s$  have been omitted.

Thus we see that the structure may produce reflexions on the meridian at both 5.17 Å and 1.48 Å. We must now estimate whether these reflexions are likely to be strong.

Consider the 5.17 Å reflexion. The amplitude is given by  $J_1(2\pi(l/c)r_1 \sin \alpha)$ , and, since  $2\pi(l/c) \sin \alpha = 0.21$ , this is  $J_1(0.21r_1)$ . Thus for the atoms of the polypeptide chain, for which  $r_1$  is between 1.6 and 2.3, the contribution to this reflexion will by itself be small. If, however, we consider the side-chain, or more plausibly the first two or three atoms of the side-chain, as a blob with centre at  $r = 4.4\frac{1}{2}$  Å, say, which is not an unreasonable postulate at this resolution, then we obtain  $J_1(0.9)$ , which has the value 0.41 and is quite considerable.

In order to give a concrete example the structure factor of this reflexion has been calculated for polyvaline. Valine has the side-chain  $-\text{CH}_2(\text{CH}_3)_2$ . The two methyl groups were placed in the most plausible positions consistent with staggering the bonds. The results are given in Fig. 5. It can be seen that the contribution  $OA$  from the polypeptide chain is nearly in phase with the side-chain contributions for the  $\beta\text{C}$  in position 2, but about 90° out of phase for position 1. The effect is that the former structure gives double the intensity of the latter.

Consider next the 1.48 reflexion. The amplitude of this is given by  $J_0(0.74r_1)$ . For  $r_1 = 1.65$  we obtain  $J_0(1.2) = 0.67$ .

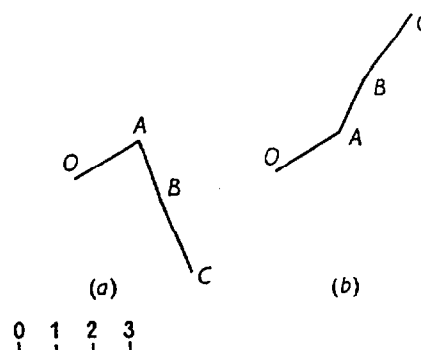


Fig. 5. The amplitude of the 5.1 Å meridional reflexion for a hypothetical coiled-coil. (a)  $\beta\text{C}$  in position 1. (b)  $\beta\text{C}$  in position 2.  $OA$ : main chain;  $AB$ :  $\beta\text{C}$ ;  $BC$ :  $(\text{CH}_3)_2$ ;  $OC$ : resultant amplitude. Scale in electrons per residue.

The exact theory shows that it is slightly less than this, but it is easy to see that the contribution from the main polypeptide chain is likely to be considerable, and qualitative comparisons with the straight  $\alpha$ -helix confirm this. The side-chains might be expected to contribute little here, since it is doubtful if they would be sufficiently ordered at this resolution.

As to the other possible meridional reflexions, we shall get very little for the 10.3 Å reflexion, since  $|s|$  is 3 and  $l$  is small. Those at 1.72 and 1.03 Å can only be weak or absent, since the main chain will contribute little, and the side-chains are unlikely to be sufficiently ordered at these resolutions. The same may be true of the 2.59 Å reflexion.

It is difficult to decide about the 3.45 and 2.07 Å reflexions without detailed calculations, but it would not be surprising if they were weak or absent, since the analogous regions in the diffraction pattern of the straight  $\alpha$ -helix are rather weak.

These results are not altered when we consider the effects of both helices together. It is easy to see from the expression for the phase in equation (1) that the contributions from the two helices related by a diad will always add if  $(p-q)$  is even, and cancel if  $(p-q)$  is odd. Examination of equation (3) shows that in this particular case it is equivalent to saying that the contributions add on even layer lines, and that odd layer lines are absent. This can be seen in another way. We have chosen our parameters so that one of the coiled  $\alpha$ -helices, considered separately, has an 18-fold

screw axis, and it will therefore have a screw diad as a component of this. Thus in this case the second half of one helix will be identical with the first half of the other, and the true repeat of the pair in the  $z$  direction will be half that of one of them considered separately. Thus odd layer lines will be absent. The contributions on the meridian from the two helices of this model will add, since we have shown that we only get non-zero intensity there when  $l$  is a multiple of 18.

Thus the coiled-coil hypothesis explains in a straightforward manner the co-existence of the 5.1 and the 1.5 Å reflexions on the meridian. The amplitude of the 1.5 Å comes from the precise repeat of the main polypeptide chain and the 5.1 Å mainly from the less precise but sufficiently regular arrangement of the side-chains, or the first two or three atoms of them. The origin of the coiled-coil is explained in an entirely natural way as the result of the close-packing of two adjacent  $\alpha$ -helices, and it is not necessary to postulate a regular sequence of residues repeating every seventh residue. The figure seven comes directly from the nature of an  $\alpha$ -helix.

We next consider the equatorial region of the diffraction pattern of the two-strand rope.

On the equator we have  $l = 0$ . Thus  $s$  must be zero, or the third Bessel term will be zero. We thus have for the structure amplitude

$$\sum_p \sum_q J_p(2\pi R r_0) J_q(2\pi R r_1),$$

subject to

$$p + 35q = 126m'.$$

For small  $R$  the only term of any importance has  $p = q = 0$ . This is the approximation used (though for their 7-core cable) by Pauling & Corey (1953). For higher  $R$  one might have to consider such terms as  $p = 7, q = 7$ ;  $p = -14, q = 4$ ;  $p = -7, q = 11$  etc. One can find such combinations systematically by plotting them on graph paper, using  $p$  and  $q$  as the two co-ordinates. The possible values will be found to form a regular pattern, and it is easy to see whether any important term has been missed.

Of the general reflexions from the two-strand rope we shall discuss only the near equatorials, i.e. those terms given by  $p \neq 0, q = s = m' = 0$ . We thus have for a single coiled coil

$$C(R, \psi, l/c) =$$

$$J_p(2\pi R r_0) J_0(2\pi R r_1) J_0(2\pi(l/c)r_1 \sin \alpha) \exp ip(\frac{1}{2}\pi + \psi),$$

subject to

$$N_0 p = l.$$

Since  $N_0 = 1$ , we have  $p = l$ . For small values of  $l$ ,  $J_0(2\pi(l/c)r_1 \sin \alpha)$  is approximately equal to 1.0. For simplicity let us consider  $r_1$  small, so that  $J_0(2\pi R r_1)$  is varying slowly with  $R$ . Then the position of the first maximum of this function will be near the first maximum of  $J_p(2\pi R r_0)$ . Taking  $r_0 = 5.2$ , we obtain

the results of Table 2 for the spacing ( $1/R$ ) of this maximum.

Table 2

$p$	1	2	3	4
$1/R$ (Å)	17.8	10.9	7.75	6.2
$\theta$ (°)	$5\frac{1}{2}$	$6\frac{1}{2}$	7	$7\frac{1}{2}$

Here  $\tan \theta = l/cR$ . The pitch angle of one coiled-coil was  $10^\circ$ . Thus the maximum of the Bessel functions of the near equatorials occurs at an angle from the equatorial direction of about two-thirds of the pitch angle. This gives us a very rough-and-ready method of estimating the pitch angle of a coiled-coil from the near equatorial region of the diffraction pattern.

As shown above, for our two-strand rope only even values of  $l$  are present, so that the first near-equatorial will be for  $l = 2$ . The maximum of  $J_2$  is fairly broad, and the exact spacing observed in practice will depend upon the distance of neighbouring parts of the structure, but we can reasonably expect a fairly strong reflexion in the 10 Å region.

The terms we have been considering correspond crudely to taking the  $\alpha$ -helix as a curved rod of electron density. Thus while we might expect the intensity for  $l = 2$  to be large, that for  $l = 4$  and higher is likely to be small, owing to the finite radius of the  $\alpha$ -helix. For example, for  $r_1 = 1.8$  Å,  $J_0(2\pi R r_1) = 0$  when  $1/R = 4\frac{1}{2}$  Å.

The remainder of the general reflexions will not be discussed here in detail. The most important terms are likely to be the 5 Å near-meridionals, which can be thought of as coming from the strong reflexion on the 5.4 Å layer line of the straight  $\alpha$ -helix, which have now been dispersed into a number of reflexions in that region of the reciprocal lattice. They are given in the theory by terms for which  $m'$  is zero,  $s$  and  $q$  are either 0 or 1, and  $p$  is small. There will also be a similar set due to the dispersal of the 1.5 Å reflexion. For these  $m' = 1, (s+q) = 0$  and  $p$  and  $q$  are small. Their main effect is likely to be the production of a broad 'arc' in that region.

### The three-strand rope

This is the next simplest case. Three  $\alpha$ -helices pack together so that they slowly coil round a straight line running between them. This structure is called the  $D_3$  rope by Pauling & Corey (1953). One possible way to relate the helices is by a triad axis. Each helix is then arranged so that the broken line in Fig. 3 always points towards this triad axis. If we make a section perpendicular to  $z$  at the appropriate level we shall obtain something like Fig. 6. Had we taken the section about 10.3 Å higher, we should have obtained an identical picture except that it had been rotated through about  $20^\circ$  about the fibre axis. If we had gone only 5.15 Å higher we should have got the mirror image of Fig. 6 reflected about the broken line, with

a  $10^\circ$  rotation superimposed. It is these layers of side-chains occurring every  $5.15 \text{ \AA}$  in the  $z$  direction which are mainly responsible for the  $5.15 \text{ \AA}$  meridional reflexion.

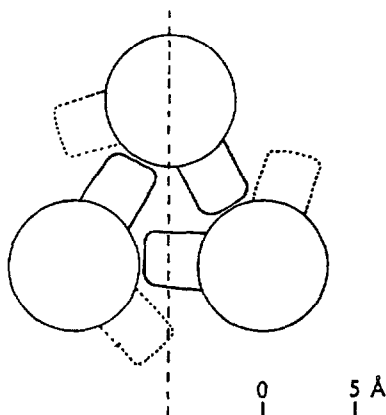


Fig. 6. Schematic diagram to illustrate the packing of the three-strand rope. The figure is a section perpendicular to the fibre axis. The circles represent the main polypeptide chains. The knobs represent side-chains. Only the inner side-chains are shown. The dotted side-chains are at a slightly different level from those drawn with solid lines.

The diffraction pattern for the three-strand rope will have broad similarities to that of the two-strand rope. We can take the values of  $N_0$ ,  $N_1$  and  $M$  to be the same as before. As each curved helix can be turned into itself by the same screw axis as before, i.e. by a translation of  $10.3 \text{ \AA}$  and a rotation of  $20^\circ$ , the meridional reflexions will occur at the same spacings. We shall have slightly larger values for  $r_0$  and  $\alpha$ , so that the intensities will not be quite the same, but the general arguments used will still apply. The intensity of the  $5.15$  reflexion is likely to be increased rather than decreased, owing to the increase in  $\alpha$ .

The fact that we have three rather than two  $\alpha$ -helices means that only every third layer line will be present, rather than every second. The fundamental repeat for this simple model will therefore be about  $186/3 = 62 \text{ \AA}$ .

#### Additional consideration

Under this heading will be discussed the effect of the slight distortions of our simple structures which prevent them repeating exactly after one turn of the major helix, the effect of repeating sequences of residues, and the effect of the helices not all running in the same direction.

Although the knobs-into-holes packing argument makes it very likely that every seventh residue will be the same distance from the fibre axis, there is no need for the major helix to rotate *exactly*  $20^\circ$  in this interval, and quite minor distortions of the helix may alter this figure appreciably. Suppose, as an illustration, that in our three-strand rope the angle was

$360^\circ/19 \approx 19^\circ$ . The structure would then repeat exactly after  $(19 \times 7) = 133$  residues. We should have  $N_0 = 1$ ,  $N_1 = 38$ ,  $m = 133$ ,  $c = 133 \times 1.5 \cos 12^\circ = 195 \text{ \AA}$  and  $r_0 = 10.4/\sqrt{3} = 6 \text{ \AA}$  (the values of  $r_0$  and  $\alpha$  are only illustrative; they will depend on the nature of the residues packing round the triad axis, and they may well be rather smaller).

We should then have, instead of equation (3), the restriction

$$p + 37q + 38s = l + 133m'.$$

The meridian reflexions will thus be sub-spacings of  $195/19 = 10.26 \text{ \AA}$ , and will be rather similar to our previous figures. (If  $r_0$  were smaller, so that  $\alpha$  were  $10^\circ$ , they would be identical with the previous values.) However, since each curved helix now has a 19-fold screw, and since 19 is prime, we shall now get layer-lines for all values of  $l$ , and not merely for  $l$  a multiple of 3. However, the rule that  $(p-q)$  must be a multiple of 3 still holds. This imposes definite restrictions on terms occurring. Thus for example the near equatorials for which  $q = s = m' = 0$ , will occur only when  $l$  is a multiple of three, as before. The main term on the equator remains unchanged, but some of the higher terms are altered.

In the general case the minor helix may not repeat an exact number of times for one term of the major helix, so that  $N_0$  will be greater than one. Because of the knobs-into-holes we might expect for these models that  $M$  would always be  $7/2$  times  $N_1$ . As an example we may take  $N_0 = 99$ ,  $N_1 = 3,800$ ,  $M = 13,300$ , so that our restriction would be

$$99p + 3,701q + 3,800s = l + 13,300m'.$$

This has two interesting consequences. First, since the true repeat is so long the strong reflexions may occur on layer lines which are not simply related. For example the  $5.1 \text{ \AA}$  reflexion will come on  $l = 3,500$ , and the first near equatorial for a three-strand rope on  $l = 297$ , giving a ratio of  $38.38/3$ , instead of  $38/3$  as in the previous case. Thus it may be impossible to index the X-ray picture in a simple and unambiguous manner. Secondly, for the equatorial reflexions it will be found that all the terms except the first can be ignored, because the order of the Bessel functions is so high. This might be expected on physical grounds since the effect of the inexact repeat will be to smear over the details of the projection in the  $z$  direction. The term for which  $p = q = s = m' = 0$  is therefore the only one that need be considered. A similar argument can be used to justify the use of this single term by Pauling & Corey (1953).

In the completely general case two helices may pack together in an unsymmetrical manner. This corresponds, in terms of Fig. 3, to drawing the broken line which represents the line of contact in some other direction than the one shown. The path this new dotted line will follow on one of our pieces of paper will not correspond to the one it follows on the other piece,

so that the relationship between the two helices will not be symmetrical. This might happen, for example, if one helix were straight and the other coiled round it as in the  $AB_2$  cable of Pauling & Corey (1953). Beyond pointing out that this may alter the ratio of the main meridian spacings (5.15 and 1.50 Å) from the value 3.50, such structures will not be discussed further here.

We next consider the effect on the diffraction pattern of a repeating sequence of residues. This has been postulated by Pauling & Corey (1953) in their model for  $\alpha$ -keratin to account for the curving of the  $\alpha$ -helices. It might also be expected if  $\alpha$ -keratin were made up of small sub-units placed end to end.\* It is easy to see that we may get extra spots, in particular on the meridian. The simplest way to allow for this complication is to add to our basic structure some extra atoms to represent a heavy side-chain. If the sequence repeats after  $Q$  residues, we consider an additional discontinuous coiled-coil, with the same parameters as the basic one, except that we now write  $M/Q$  instead of  $M$ , since our extra atom occurs only that number of times.

Thus for our simple two-strand rope if we postulate every 12th residue the same, so that  $M/Q = 126/12 = 21/2$ , we should obtain for meridional reflexions the condition

$$36s = l + (21/2)m'$$

instead of

$$36s = l + 126m'.$$

This implies that the structure now only repeats exactly after twice the distance it did before, so that if we put  $l' = 2l$ , where  $l'$  is the index appropriated to the longer repeat, we have

$$72s = l' + 21m'.$$

Thus not only do we get the obvious extra meridional reflexions, given in our terminology by  $s = 0$ , but also other spots as well. However, we should expect the structure factors for such spots to be small, since not many atoms would be contributing to them. It should be remembered, however, that the Lorentz factor for meridional spots often makes them appear stronger on the actual X-ray photograph than reflexions in general positions.

It should be noticed that if  $l$  is 7, we do not obtain extra spots on the meridian, since on our hypothesis  $M$  will necessarily be a multiple of 7. The effect of such a repeating sequence on meridional reflexions is merely to alter the intensities of the reflexions already discussed.

In our considerations so far we have assumed that the chains all run in the same direction, so that they can be related by simple diads or triads. If they run in different directions we may expect the relationship between the side-chains that pack together to remain roughly the same. Thus the 5.1 Å meridian reflexion

for these simple models is likely to remain substantially unaltered. Moreover, the 10 Å equatorial and near-equatorial reflexions should be relatively unaffected, since at this resolution the direction of the polypeptide chain makes little difference. The 1.5 meridional reflexion, however, which comes mainly from the atoms of the backbone of the helix, will depend on the exact phasing, in the  $z$  direction, of the backbones of chains running in opposite directions. Its intensity could only be decided from a detailed examination of models.

### Possible applications

#### *Tropomyosin*

The simple structures described so far—the two-strand rope and the three-strand rope—may be the basis of long thin molecules, such as tropomyosin, which are only two or three polypeptide chains thick (Bailey, 1953). The X-ray diagram to be expected from a loose array of such molecules would have the broad characteristics already described i.e. a diffuse  $\alpha$  diagram. One is unlikely to detect very long spacings, unless the orientation and packing were particularly good, since the only strong reflexions to show a long axial periodicity are the near-equatorials, and these would probably be merged with the true equatorials. We should expect a diffuse meridional arc in the 5 Å region, with an emphasis on the meridian at about 5.15 Å. The equatorial reflexions near 10 Å would be expected to be rather diffuse and perhaps at a rather greater average spacing than usual, owing to the difficulties of packing. This is a fair description of the X-ray results of Astbury, Reed & Spark (1948) for stretched films of tropomyosin. These films also give a 1.5 Å meridional reflexion (Astbury, reported by Perutz, 1952), as one would predict.

One interesting possibility arises from models of this kind. It is noticeable that the proportion of non-polar to polar groups in tropomyosin is lower than for a more globular protein like haemoglobin. It is plausible to postulate that these non-polar groups tend to occur between the chains, rather than on the outside of the molecule, since the packing of hydrophobic groups together would allow more hydrogen bonds to be made elsewhere, and thus is energetically preferable. It is noteworthy that no straight hydrocarbon side-chains (beyond alanine) occur in proteins. The more awkward, less flexible nature of the valine, leucine and isoleucine may make it easier for the  $\alpha$ -helices to fall into a coiled-coil configuration. One might expect these non-polar side-groups to make up about two-sevenths of the residues. It is probably a coincidence, but the analysis of tropomyosin by Bailey (1948) shows that alanine plus valine plus the leucines make up 0.29 of the total number of residues, which is very close to two-sevenths. If this hypothesis were correct, one would expect there to be a tendency towards an alteration of groups of polar and non-polar residues in tropomyosin, the non-polar occurring at an average

\* See Farrant, Rees & Mercer, 1947.



interval of  $3\frac{1}{2}$  residues, so that they always pointed 'inside' the molecule. The experimental difficulties of determining the sequence or partial sequence of very long polypeptide chains are at the moment very great, but the possibility of obtaining a result of this sort should not be overlooked.

#### $\alpha$ -Keratin

A glance at the X-ray diagram of African porcupine quill (McArthur, 1943), which gives the most perfect  $\alpha$ -keratin pictures, shows that we have to deal with a structure of great complexity. It is not proposed to discuss it here in detail. The following remarks should be regarded as preliminary.

At first sight the protein appears to be a very good example of the three-strand rope (the  $D_3$  rope of Pauling & Corey, 1953). The repeat is 198 Å and, as far as can be seen from McArthur's Fig. 2, the near equatorials occur at the right place. However, the three-strand rope would not easily explain the equatorial reflexion at 27 Å, nor the high density (about  $1.3 \text{ g.cm.}^{-3}$ ).

The model proposed by Pauling & Corey (1953) partly overcomes these objections. It remains to be seen whether this model can be generated by packing considerations rather than by repeating sequences of residues, as postulated by Pauling & Corey.

While this model appears to explain the very strong reflexions, it is unlikely in its simple form to predict the rich spectrum of reflexions on or very close to the meridian. It is possible that these are produced because the structure is made out of sub-units, joined end to end, or perhaps because of distortions due to some complicated interaction between chains.

#### Globular proteins

The implication for globular proteins, if they are mainly made up of  $\alpha$ -helices, is that the  $\alpha$ -helices may well pack some  $20^\circ$  or so away from parallel. This may go a long way to explain the difficulties experienced in attempting to interpret the diffraction patterns of protein crystals. Riley & Arndt (1953) make a similar point when mentioning their attempts to build models of insulin, and it will be interesting to see if the skew structures they mention are due to an approximation to knobs-into-holes packing, or to some other reason.

The curious way that one  $\alpha$ -helix can roll round another while preserving this general sort of pack tempts one to speculate whether this may occur in enzyme action, since it would provide a mechanism whereby two small molecules, each rather rigidly attached to part of the protein surface, could be brought into proximity in a specific manner. It is clear that these questions are unlikely to be settled until

the architecture of some of the globular proteins is definitely established, and this remains the major aim of protein crystallography.

#### Conclusion

The main conclusion of this paper is that it is no longer possible to claim that the  $\alpha$ -keratin pattern cannot be explained in terms of the  $\alpha$ -helix. The general features of the observed X-ray pattern can be predicted merely by postulating that  $\alpha$ -helices tend to pack side-by-side in a knobs-into-holes manner. It is not necessary to postulate a repeating sequence of residues. The inherent plausibility of such a form of packing, and the entirely unforced way in which it leads to the main characteristics of the observed X-ray diagrams, make it probable that it forms the basis of the structure of the fibrous proteins.

*Note added in proof, 22 July 1953.*—Astbury & Hagghith (1953) have recently reported meridional reflexions for  $\alpha$ -keratin at spacings of  $3.4_0$  and  $2.5_8$  Å; intensities ( $w-m$ ).

I should like to thank Sir Lawrence Bragg and also my colleagues in the Medical Research Council Unit both for the interest they have shown in this problem and for helpful suggestions and criticisms.

#### References

- ASTBURY, W. T., REED, R. & SPARK, L. C. (1948). *Biochem. J.* **43**, 282.
- ASTBURY, W. T. & HAGGITH, J. W. (1953). *Biochim. biophys. Acta*, **10**, 483.
- BAILEY, K. (1948). *Biochem. J.* **43**, 271.
- BAILEY, K. (1953). *Proc. Roy. Soc. B*, **141**, 45.
- COCHRAN, W., CRICK, F. H. C. & VAND, V. (1952). *Acta Cryst.* **5**, 581.
- CRICK, F. H. C. (1952). *Nature, Lond.* **170**, 882.
- CRICK, F. H. C. (1953). *Acta Cryst.* **6**, 685.
- FARRANT, J. L., REES, A. L. G. & MERCER, E. H. (1947). *Nature, Lond.* **159**, 535.
- MCCARTHER, I. (1943). *Nature, Lond.* **152**, 38.
- PAULING, L. & COREY, R. B. (1951). *Proc. Nat. Acad. Sci., Wash.* **37**, 261.
- PAULING, L. & COREY, R. B. (1953). *Nature, Lond.* **171**, 59.
- PAULING, L., COREY, R. B. & BRANSON, H. R. (1951). *Proc. Nat. Acad. Sci., Wash.* **37**, 205.
- PERUTZ, M. F. (1951). *Nature, Lond.* **167**, 1053.
- PERUTZ, M. F. (1952). *Rep. Progr. Chem. for 1951*, **48**, 371.
- RILEY, D. P. & ARNDT, U. W. (1952). *Nature, Lond.* **169**, 138.
- RILEY, D. P. & ARNDT, U. W. (1953). *Nature, Lond.* **171**, 144.
- YAKEL, H. L., PAULING, L. & COREY, R. B. (1952). *Nature, Lond.* **169**, 920.